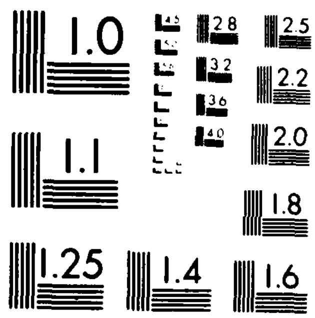
AD-A135 792 ON THE ROBUST RANK ANALYSIS OF LINEAR MODELS WITH NONSYMETRIC ERROR DISTR..(U) WESTERN MICHIGAN UNIV KALAMAZOO DEPT OF MATHEMATICS G L SIEVERS ET AL. 1/1

UNCLASSIFIED SEP 83 TR-70 N00014-78-C-0637 F/G 12/1 NL

END
DATE
FILMED
1 84
DTIC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

(16)

DEPARTMENT OF MATHEMATICS
WESTERN MICHIGAN UNIVERSITY
Kalamazoo, Michigan
49008

ON THE ROBUST RANK ANALYSIS OF LINEAR MODELS WITH NONSYMMETRIC ERROR DISTRIBUTIONS

by

Gerald L. Sievers and Joseph W. McKean

TECHNICAL REPORT NO. 70
September 1983

This research was partially supported by the Office of Naval Research under contract N00014-78-C-0637 (NR 042-407)

DEPARTMENT OF MATHEMATICS
WESTERN MICHIGAN UNIVERSITY
KALAMAZOO, MICHIGAN 49008

Summary

The robust analysis of linear models based on R-estimates involves an estimate of a scale parameter which is used in the analysis as a standardizing constant. The consistency of previous estimates of this scale parameter required that the underlying errors be symmetrically distributed. This assumption is not always warranted, for instance in survival models. A new estimate is proposed for the scale parameter and it is shown to be consistent for nonsymmetric and symmetric error distributions. With this new scale estimate, a complete robust analysis of a linear model can be accomplished without assuming symmetry. The small sample properties of the analysis are examined in a Monte Carlo study of several different situations.

Some keywords: general linear hypothesis; linear model; R-estimates; rank statistics; regression; robust analysis.

A1

1. Introduction and Summary.

McKean and Hettmansperger (1976, 1978) proposed a robust analysis of linear models based on Jaeckel's (1972) robust R-estimate of regression coefficients. This analysis is analogous to the least squares analysis of variance, indeed the geometry of the two analyses are quite similar, McKean and Schrader (1980). The R-analysis is also a generalization of R-procedures in the simpler location problems and, in particular, it inherits their asymptotic relative efficiency properties with respect to least squares. Small sample power studies, Section 4 and Hettmansperger and McKean (1983), support these asymptotic efficiencies. The R-analysis, thus, offers the user a versatile, robust alternative to least squares for analyzing linear models.

The R-analysis requires the estimation of a scale parameter, τ, which is used as a standardizing factor in the associated R-test for a general linear hypothesis and in the standard error of an R-estimate. The estimate of τ proposed by McKean and Hettmansperger (1976) is consistent under the assumption of symmetrically distributed errors. There are many situations, though, when the assumption of symmetry is unrealistic. For example many of the parametric models used for survival time have skewed distributions, see Chapter 2 of Miller (1981). In accelerated failure time models, Chapter 2 of Kalbfleisch and Prentice (1980), the log of survivor time is cast in a linear model whose error structures have skewed distributions. Many of these skewed distributions also have long tails in which case a robust analysis is especially attractive.

In this paper we propose an estimate of τ for the case of Wilcoxon scores. The proof of its consistency, Sections 3 and 5, does not require symmetry. Using this estimate of scale the R-analysis of McKean and

Hettmansperger is asymptotically valid for both symmetric and non-symmetric error distributions. Further, as noted in Section 2, by utilizing the theory of Sievers (1983) the asymptotic theory of the R-analysis based on Wilcoxon scores can be obtained under much milder regularity conditions on the design matrix than those assumed by Jureckova (1971) and Jaeckel (1972).

Asymptotic theory serves as a useful guideline for the R-analysis but for practical use an investigation of its small sample properties is required as discussed in Section 4. The estimate of τ is based on residuals and certain small sample corrections are necessary for its use in standardizing the R-test statistics. We present the results, Section 4, of a Monte Carlo study which included both symmetric and skewed distributions over several different designs. On the basis of this study, the R-analysis utilizing the new estimate of τ performed as well as that using the old for the symmetric distributions and performed better for the skewed distributions. Both the old and the new were robust with respect to least squares. On the basis of the theory and this study, we would recommend using the R-analysis with the estimate of τ found in Section 3.

2.1 Notation and assumptions.

Let Y be an $n \times 1$ vector of observations which follows the linear model

$$Y = \alpha 1 + X\beta + e \qquad (2.1)$$

where α is the intercept parameter, 1 denotes an $n \times 1$ vector of ones, β is a $p \times 1$ vector of parameters, X is an $n \times p$ design matrix, and e is an $n \times 1$ vector of independent and identically distributed errors which have the common density function $f(x)$. We will assume that X has full

column rank p and, since the model includes an intercept parameter, that the column averages of X are zero. We denote general linear hypotheses as,

$$H_0: H\beta = 0 \qquad \text{versus} \qquad H_A: H\beta \neq 0, \tag{2.2}$$

where H is a $q \times p$ matrix of full row rank.

Let Ω denote the column space of X. We can express the model (2.1) equivalently as

$$Y = \mu + e^* , \quad \mu \varepsilon \Omega \tag{2.3}$$

where $e^* = \alpha 1 + e$. Letting ω denote the $(p - q)$-dimensional subspace of Ω constrained by $H\beta = 0$, the hypotheses (2.2) can then be expressed as

$$H_0: \mu \varepsilon \omega \qquad \text{versus} \qquad H_A: \mu \varepsilon \Omega\backslash\omega .$$

Let P_Ω denote the projection matrix onto Ω.

The asymptotic distribution theory discussed in this paper requires some assumptions on the subspace Ω and the underlying density $f(x)$. We will consider sequences of subspaces Ω_n, indexed by the sample size n, which have a common dimension p and which satisfy

$$\lim_{n \to \infty} \max_{1 \leq i \leq n} h_{iin} = 0 \tag{A.1}$$

where h_{iin} is the ith diagonal element of the projection matrix P_{Ω_n}. For asymptotic inference concerning β we will further assume that the sequence of design matrices X_n satisfies

$$\lim \frac{1}{n} X'_n X_n = \Sigma, \ \Sigma \text{ is positive definite.} \tag{A.2}$$

We will often suppress the subscript n on Ω and X.

For the density function $f(x)$ we will assume the following:

$$\begin{aligned}
&\text{(i)} \quad f \text{ is absolutely continuous;} \\
&\text{(ii)} \quad \int (f'/f)^2 f \, dx < \infty; && \text{(B.1)} \\
&\text{(iii)} \quad \tau = [12^{1/2} \int f^2 dx]^{-1} < \infty.
\end{aligned}$$

2.2 R-analysis.

The part of model (2.1) involving β is often termed the deviations from means model. Least squares inference for β is based on the squared semi-norm

$$||u||_{LS}^2 = \sum (u_i - \bar{u})^2, \; u \in R^n, \tag{2.4}$$

which can also be written as

$$||u||_{LS}^2 = n^{-1} \sum_{i<j} \sum (u_i - u_j)^2 \tag{2.5}$$

For the Wilcoxon inference on β, the squared differences in (2.5) are replaced by the absolute differences; that is, the R-semi-norm is given by

$$||u||_R = (2(n+1)/\sqrt{12}) \sum_{i<j} \sum |u_i - u_j|, \; u \in R^n. \tag{2.6}$$

The factor in parentheses is a standardization constant. The R-semi-norm can be expressed more familiarly as

$$||u||_R = (\sqrt{12}/(n+1)) \sum (R(u_i) - ((n+1)/2)) u_i,$$

where $R(u_i)$ is the rank of u_i among $u_i, \ldots, u_n$, see Hettmansperger and McKean (1978).

The estimate of β obtained from this semi-norm, that is, a value of β which minimizes $||Y - X\beta||_R$, is the R-estimate proposed by Jaeckel (1972) for Wilcoxon scores. In terms of Model (2.3), McKean and Schrader (1980) defined $\hat{\mu}_R$ as a best R-predictor of Y when $\hat{\mu}_R$ minimizes

$||Y - \mu||_R$, $\mu \in \Omega$. Jaeckel obtained the asymptotic distribution of these R-estimates under conditions more restrictive than those above; however, Sievers (1983) showed that under (A.1), (A.2), and (B.1) the distribution of $\hat{\beta}_R$ is approximately $N(\beta, \tau^2(X'X)^{-1})$ where τ is defined in (B.1).

The least squares test statistic of the hypotheses (2.2) is based on a comparison of the squared distance, in terms of the norm (2.4), between Y and each of the subspaces ω and Ω. Analogously McKean and Hettmansperger (1976) proposed as an R-test statistic

$$F_R = \frac{||Y - \hat{\mu}_\omega||_R - ||Y - \hat{\mu}_\Omega||_R}{(p - q)\hat{\tau}/2} \tag{2.8}$$

where $\hat{\mu}_\omega$ and $\hat{\mu}_\Omega$ are best R-predictors of Y from the subspaces ω and Ω and $\hat{\tau}$ is an estimate of τ. Similar to least squares, the numerator of F_R is a comparison of R-distances between Y and the subspaces ω and Ω. The test is to reject H_0 in favor of H_A for large values of F_R. McKean and Hettmansperger showed that under H_0 and regularity conditions, $(p - q)F_R$ has an asymptotic chi-squared distribution with $(p - q)$ degrees of freedom, provided $\hat{\tau}$ is a consistent estimate of τ. Using Sievers' (1983) development, the conditions (A.1), (A.2), and (B.1) suffice for the regularity conditions.

A Wald type of test for H_0, based on the full model R-estimate, is to reject H_0 for large values of the statistic,

$$B = \frac{(H\hat{\beta}_R)'(H(X'X)^{-1}H')^{-1}(H\hat{\beta}_R)}{(p-q)\hat{\tau}^2} \tag{2.9}$$

Unlike their least squares counterparts the R-statistics B and F_R are not algebraically equal; however they have the same asymptotic null distribution and asymptotic relative efficiency, see Hettmansperger and McKean (1983). The statistic B is similar to Bickel's (1976) test based upon pseudo-observations obtained from an M-estimate of β.

Although the asymptotic theory suggests chi-squared critical values, from previous Monte Carlo studies we have found it best to compare the statistics F_R and B with F-critical values having $p - q$ and $n - p$ degrees of freedom. This is discussed further in Section 4.

3. Estimate of τ.

In order to use the R-analysis, the scale parameter τ, (B.1), needs to be estimated. In the simple location problem under the additional assumption of a symmetric distribution, Lehmann (1963) proposed a consistent estimate of τ based on the length of a distribution free confidence interval for the location parameter. For the R-analysis of linear models, McKean and Hettmansperger (1976, 1978) proposed using Lehmann's estimate of τ computed from the residuals. Under the assumption of symmetrically distributed errors they showed the estimate was consistent. For the asymptotic distribution theory of the R-analysis, though, this was the only place where symmetry was invoked. In this section we propose a new estimate of τ and show that it is consistent under the assumptions (A.1), (A.2), and (B.1). The R-analysis of Section 2.2 can thus be employed for both symmetrically and nonsymmetrically distributed errors.

We find it more convenient to work with the parameter γ defined by

$$\gamma = (\sqrt{12}\ \tau)^{-1} = \int f^2 \tag{3.1}$$

Note that the difference of a pair of error random variables, $e_i^* - e_j^* = e_i - e_j$, has as its density,

$$g(t) = \int f(x) f(x - t) dx. \tag{3.2}$$

This suggests estimating γ with an estimate of $g(0)$.

Let Z denote the vector of R-residuals,

$$Z = Y - \hat{\mu}_R \tag{3.3}$$

where $\hat{\mu}_R = X\hat{\beta}_R$ is a best R-predictor of Y from Ω. We are interested in the differences of the residuals, $Z_i - Z_j$, and since these are symmetric about 0, we need only consider the set of $|Z_i - Z_j|$ for $i < j$. The corresponding empirical distribution function is

$$\hat{G}_n(t) = \binom{n}{2}^{-1} \sum_{i<j} \phi(|Z_j - Z_i|, t) \tag{3.4}$$

where $\phi(u,v) = 1$ or 0 if $u \leq v$ or $u > v$. Our estimate of γ is a measure of the slope of $\hat{G}_n$ at 0, namely

$$\hat{\gamma} = \hat{G}_n(\hat{t}_\alpha/\sqrt{n})/(2\hat{t}_\alpha/\sqrt{n}) \tag{3.5}$$

where $\hat{t}_\alpha$ is the αth quantile of $\hat{G}_n(t)$. The corresponding estimate of τ is

$$\hat{\tau} = (\sqrt{12}\,\hat{\gamma})^{-1}. \tag{3.6}$$

Some discussion on the choice of α can be found in Section 4.

The consistency of $\hat{\tau}$ is noted in the following theorem which is proved in more generality in Section 5.

<u>Theorem 3.1</u>. Consider a sequence of linear models which satisfies assumptions (A.1), (A.2), and (B.1). For $0 < \alpha < 1$, $\hat{\tau} \xrightarrow{P} \tau$.

If τ is estimated by $\hat{\tau}$ of (3.6) then the R-analysis discussed in Section 2.2 is asymptotically valid for both symmetric and non-symmetric error distributions. It can thus be used for instance in the applications mentioned in Section 1. The small sample properties of $\hat{\tau}$ are discussed in the next section.

The algorithm we employ to compute $\hat{\gamma}$ is an iterative procedure which exploits the relationship between $\hat{G}_n$ and the Wilcoxon two sample process defined below. This procedure avoids the storage and sorting of the $\binom{n}{2}$ differences, requiring only vectors of length $2n$.

In order to define the two sample process consider as one sample the residuals $Z_1,\ldots,Z_n$ and as a second sample exactly the same set of residuals. The Wilcoxon process is

$$W_n(t) = \sum_{i=1}^{n} R(Z_i - t) - n(n + 1)/2$$

where $R(Z_i - t)$ denotes the rank of $Z_i - t$ among $Z_1 - t,\ldots,Z_n - t,\ Z_1,\ldots,Z_n$. For $t > 0$, it can be sho ɩnat

$$\hat{G}_n(t) = 1 - \binom{n}{2}^{-1} W_n(t).$$

The quantile t_α can be obtained by solving the equation $1 - \binom{n}{2}^{-1} W_n(t_\alpha) = \alpha$, which can be solved iteratively using a modified version of regular falsi similar to the algorithm presented by McKean and Ryan (1977).

In the simple location problem Schweder (1975) proposed an estimate of γ which is basically a window estimate and he recommended a uniform window symmetric about zero. Assuming $\mu_i = 0$ for all i so that no nuisance parameters need be estimated, his estimate is

$$\begin{aligned} &\sum_{i,j} \phi(|Y_j - Y_i|,\ h_n/2)/n^2 h_n \\ &= ((n - 1)/n) G_n(h_n/2)/h_n + (1/nh_n) \\ &\doteq G_n(h_n/2)/h_n \end{aligned}$$

where h_n is the width of a uniform window. The estimate (3.5) is seen to be similar to this type of estimate with a window width of $O_p(n^{-1/2})$. In proving consistency and asymptotic normality, Schweder assumed the window width satisfies $\sqrt{n}\, h_n \to \infty$, ruling out the $O(n^{-1/2})$ cases, and recommended

$h_n = O(n^{-1/3})$.

Our estimate $\hat{\gamma}$ differs from Schweder's estimate by allowing for the estimation of nuisance parameters in using residuals rather than iid variables. Also it uses a window width estimated from the data. In an actual application of Schweder's estimate the choice of an appropriate s for $h_n = sn^{-1/3}$ is difficult to make and his Monte Carlo study indicates that the bias of the estimate is sensitive to the choice made.

4. Small Sample Properties

In this section we discuss a Monte Carlo study of the performance of $\hat{\tau}$ on F_R and B in several small sample cases that reflect problems arising in practice. The results show that some modification is necessary to obtain satisfactory results. After some small sample corrections are made, the new R-method performs quite well. It is a definite improvement over the old R-method for nonsymmetric distributions and is superior to least squares methods in maintaining level and power for nonnormal distributions.

The performance of $\hat{\tau}$ as an estimate of τ can be studied by examining bias and mean square error. Instead of this we chose to examine directly how well $\hat{\tau}$ standardizes the test statistics F_R and B. The inclusion of B in the study gives information on how well $\hat{\tau}$ performs for confidence intervals since confidence regions for β based on $\hat{\beta}_R$ can be obtained by inverting the test statistic B for appropriate H matrices.

The scale estimate $\hat{\tau}$ is based on residuals which tend to be less variable than independent observations. This effect diminishes for large sample sizes, but for smaller sample sizes the dependence among residuals may cause τ to regularly underestimate τ. An indication of this effect is shown in Figure 4.1. This figure shows graphs of the empirical distribution functions

functions of the differences $|e_i - e_j|$ of random errors and of the differences of R-residuals $|Z_i - Z_j|$ for one simulation of lognormal variables in design A below. This figure is typical of many others we have examined. Note that the differences of R-residuals are stochastically smaller than the differences of the independent observations. As a result, the estimate of slope $\hat{\gamma}$ is larger for residuals than for actual errors and the reciprocal is then biased downward.

One of the more successful corrections we considered was the standard least squares correction, namely,

$$\hat{\tau}^* = \hat{\tau}(n/(n - p))^{1/2}. \quad (4.1)$$

As noted in Figure 4.1, this has the effect of increasing the horizontal coordinate used in the slope estimate and thus decreasing the slope.

Another small sample correction that improved the empirical results is to use a critical value from the F distribution with $p - q$ and $n - p$ degrees of freedom instead of a chi-square critical point. The R-analysis then uses the same critical values as the least squares analysis.

4.2 Monte Carlo Study

The empirical study here considers the tests based on F_R and B with τ^* of (4.1) used in the denominator. Results are reported here for τ^* based on quantiles $\alpha = .80$ and $.90$ with the corresponding test statistics denoted $F_R(\alpha)$ and $B(\alpha)$. The rank statistic of Hettmansperger and McKean (1976, 1978), which is appropriate for symmetric error distributions, is also included and denoted by F_0. It uses the numerator of F_R standardized by Lehmann's estimate of τ (see section 3). The specific estimate we used is given in formula (2.7) of Hettmansperger and McKean (1978) based on an 80%

confidence level and with the same small sample correction as expressed in (4.1) here. The least squares analysis, denoted by LS, is also included to provide a reference point and show the potential gain in using the R-analysis instead of the classical analysis.

The R-analyses were computed by using the k-step R-estimates with $k = 2$. McKean and Hettmansperger (1978) noted the close agreement between the analyses based on the 2-step estimates and those based on the fully iterated estimates, which we have confirmed in a small pilot study. The analyses based on the k-step estimates have the same asymptotic properties while taking much less time to compute. The M-estimates of Huber (1973) were used as starting values. The estimates of τ were computed by an iterative routine as noted in Section 3. The actual algorithms used are described in Hettmansperger and McKean (1983).

We examined the behavior of the tests over three designs. One was the same unbalanced 2 x 2 design considered by McKean and Hettmansperger (1978) which has cell sizes of $n_{11} = n_{22} = 8$ and $n_{12} = n_{21} = 5$. We considered the three hypotheses: H_{01}, average row effects the same; H_{02}, average column effects the same; and H_{03}, no iteraction. For this design, which we label A, the sample size is 26, $p = 4$, and $p - q = 1$ for all three hypotheses.

The other two designs each involve two regression lines. The design matrices are of the form,

$$X = \begin{pmatrix} 1 & 0 & x & 0 \\ 1 & 1 & x & x \end{pmatrix} .$$

For design B the 1 denotes a vector of 10 ones and x is the vector $(-1,-.8,\ldots,-.2,.2,\ldots,1)'$. Design C was a larger version of B, with the x vector consisting of the 20 values, $(-1, -.9,\ldots,.1, .1,\ldots,1)'$. For each design we considered the three hypotheses: H_{01}, same intercept; H_{02},

same slope; and H_{03}, same regression line. Thus the sample sizes for the designs are 20 and 40, p is 4 for both designs, and p - q is 1 for the first two hypotheses and 2 for the third hypothesis.

Four error distributions were considered in this study. Two of the distributions were symmetric: the standard normal (N) with distribution function $\Phi(x)$ and a contaminated normal (CN) with distribution function $F(x) = .85\,\Phi(x) + .15\,\Phi(x/7)$. The other two distributions were skewed: a lognormal (L) with distribution function $F(x) = \Phi(\log x)$ and a skewed contaminated normal (SCN) with distribution function $F(x) = .82\,\Phi(x) + .18\Phi((x - 1.9)/13.5)$. The last is a moderately skewed distribution proposed by Draper (1981).

The normal variates were obtained by the transformation on a pair of uniforms as proposed by Marsaglia and Bray (1964) while the uniforms were generated by the algorithm UNI developed by Gross (1979).

For each situation, a design and a distribution, we ran 1000 simulations. Both empirical levels and powers were investigated. We chose alternatives to give, for the most part, a spectrum of the power curves.

4.3. Results.

We will summarize our results for the tests in terms of empirical levels and powers for three alternatives. For a nominal .05 level, the empirical levels are recorded in Table 4.1 and the powers are in Tables 4.2 - 4.5. The results for nominal levels .01 and .10 were similar. Since the results are based on 1000 simulations, two standard errors for the levels is about .014. In Table 4.1 a plus indicates that the value is above .05 by two standard errors while a minus indicates that it is below.

The effect of α on the estimate of τ is apparent from Table 4.1. The levels for $\alpha = .90$ are generally much more conservative than those for $\alpha = .80$. Although not tabled, we found that this trend continued for $\alpha = .95$. The choice of α seems important. In this study the value of $\alpha = .80$ produced the best results. Note that the Wald type test is slightly more liberal, at $\alpha = .80$, than the drop in dispersion test. The comparison of interest is between D_0 and $F_R(.8)$. Their empirical levels for the symmetric distributions are quite close. This is true for the skewed contaminated normal; however, in the case of the lognormal the levels of D_0 are much smaller.

For the empirical powers we only tabled the results of the procedures based on $\alpha = .80$. The table index for the alternatives is the parameter $\lambda = 2^{-1}(H\beta)'(H(X'X)^{-1}H')^{-1}H\beta$ where β is the vector of alternatives. Hence λ is the non-centrality parameter for the least squares F-test at the standard normal.

For the normal distribution, least squares dominates, as it theoretically should, over all the R-procedures but not by that much. The R-procedures essentially have the same power, among themselves. The reverse is true for the contaminated normals where all the R-procedures are substantially more powerful than least squares over all the designs. Among the R-procedures, perhaps D_0 has a slight edge. The same is true for the skewed contaminated normals except that the R-procedures based on the new estimate of τ behave slightly better. For the lognormal distribution the new R-methods are more powerful than D_0 over all the designs. In this case the procedure D_0 tends to be conservative, especially in the large Design C; however, even here it is more powerful than least squares.

5. Consistency for the Scale Estimate

We will consider a more general estimate of τ based upon a weighted empirical distribution function given by

$$\hat{H}_n(t) = \sum_{i<j} a_{ij}\phi(|Z_i - Z_j|, t) \tag{5.1}$$

where Z is the vector of R-residuals given by (3.3) and the weights a_{ij} are non-negative, $\Sigma a_{ij} = 1$, and satisfy

$$n^2 \sum a_{ij}^2 \leq B_0^2 \tag{5.2}$$

for some constant B_0. The unweighted case, $a_{ij} = \binom{n}{2}^{-1}$, provides the estimate of τ found in Section 3. Other weights can be used for special purposes. For instance, in analysis of variance models and stratified sampling models where the data arise naturally in groups, the use of weights $a_{ij} = 0$ if Z_i, Z_j are from different groups would leave the estimate dependent only on the within-group variation of the data. Estimates with this feature have been proposed by Draper (1981). As in Section 3, for $0 < \alpha < 1$ define an estimate of γ by

$$\hat{\gamma}_{n\alpha} = \hat{H}_n(\hat{t}_\alpha/\sqrt{n})/(2\,\hat{t}_\alpha/\sqrt{n}) \tag{5.3}$$

where $\hat{t}_\alpha$ is the αth quantile of $\hat{H}_n(t)$.

Consider a sequence of models $Y_n = \mu_n + e_n$ where $\mu_n \in \Omega_n$, Ω_n satisfies the assumptions (A.1) and (A.2), and the vector of errors e_n satisfies (B.1). Since the estimator $\hat{\gamma}_n$ is translation invariant, without loss of generality we will assume that the true value of μ_n is 0. These assumptions will hold for the remainder of the paper.

In order to facilitate the following proofs consider the subset M_n of

Ω_n defined by

$$M_n = \{\mu \in \Omega_n : ||\mu|| \leq \Delta\}, \tag{5.4}$$

where $\Delta > 0$ is a given constant. By assumption (A.1) it follows that,

$$\lim_{n\to\infty} \sup_{\mu_n \in M_n} \max_{1\leq i<j\leq n} |\mu_{jn} - \mu_{in}| = 0. \tag{5.5}$$

For $\delta > 0$ we construct a covering $M_{1n}, \ldots, M_{Kn}$ of M_n as follows: Let C_n be an $n \times p$ matrix whose columns form an orthonormal basis for Ω_n. Let $D = \{\theta \in R^p : ||\theta|| \leq \Delta\}$, where Δ is given in (5.4). Since D is compact we can find an integer K along with points $\theta_1, \ldots, \theta_K$ in D such that $D \subset \bigcup_k D_k$ where $D_k = \{\theta : ||\theta - \theta_k|| \leq \delta\}$. Defining $M_{kn} = C_n(D_k)$ for $k = 1, \ldots, K$, it follows that $M_n \subset \bigcup_k M_{kn}$. For any such covering of M_n let

$$\begin{aligned} d^U_{ijkn} &= \sup\{\mu_{jn} - \mu_{in} : \mu_n \in M_{kn}\} \\ d^L_{ijkn} &= \inf\{\mu_{jn} - \mu_{in} : \mu_n \in M_{kn}\}, \end{aligned} \tag{5.6}$$

where $k = 1, \ldots, K$ and $1 \leq i < j \leq n$. We need the following inequality:

<u>Lemma 5.1</u>. For $\delta > 0$ let $M_{1n}, \ldots, M_{Kn}$ be the associated covering of M_n. Then for $k = 1, \ldots, K$

$$n^{-1} \sum_{i<j} (d^U_{ijkn} - d^L_{ijkn})^2 \leq 2p\delta^2 .$$

<u>Proof</u>:

Let c_{in} denote the ith row of the matrix C_n. Then $d^U_{ijkn} = \sup\{(c_{jn} - c_{in})'\theta : \theta \in D_k\}$. Since D_k is compact we have $d_{ijkn} = (c_{jn} - c_{in})'\theta^U_{ijk}$ for some $\theta^U_{ijk} \in D_k$. Similarly $d^L_{ijkn} = (c_{jn} - c_{in})'\theta^L_{ijk}$ for some $\theta^L_{ijk} \in D_k$. By the Cauchy-Schwarz inequality we then obtain

$$(1/n)\sum_{i<j}(d^U_{ijkn} - d^L_{ijkn})^2 \leq (1/n)\sum_{i<j}||c_{jn} - c_{in}||^2||\theta^U_{ijk} - \theta^L_{ijk}||^2$$
$$\leq 2\delta^2 \sum_{\ell=1}^{p}(1/n)\sum_{i<j}(c_{j\ell n} - c_{i\ell n})^2 = 2\delta^2 p.$$

Define the process

$$R_n(t, \eta_n, Y_n) = \sqrt{n}\, H_n(t/\sqrt{n}, \eta_n, Y_n) - 2t\gamma,$$

where $H_n(t/\sqrt{n}, \eta_n, Y_n) = \sum_{i<j} a_{ij}\, \phi(|(Y_{jn} - \eta_{jn}) - (Y_{in} - \eta_{in})|, t/\sqrt{n})$.

<u>Lemma 5.2</u>. Under the above assumptions, for $\eta_n \in M_n$ and $t > 0$, $R_n(t, \eta_n, Y_n) \xrightarrow{P} 0$ as $n \to \infty$.

<u>Proof</u>:

$$R_n(t,\eta_n,Y_n) = \sqrt{n}\sum_{i<j} a_{ij}(W_{ij} - 2t\gamma/\sqrt{n})$$

where $W_{ij} = \phi(|Y_j - Y_i - d_{ij}|, t/\sqrt{n})$ and $d_{ij} = \eta_{jn} - \eta_{in}$. Using

$$E(W_{ij}) = G(d_{ij} + t/\sqrt{n}) - G(d_{ij} - t/\sqrt{n})$$
$$= g(\xi_{ij})(2t/\sqrt{n})$$

where $|\xi_{ij} - d_{ij}| < t/\sqrt{n}$ and $\gamma = g(0)$, it follows that

$$E(R_n(t,\eta_n,Y_n)) = 2t\sum_{i<j} a_{ij}(g(\xi_{ij}) - g(0)).$$

By (5.2), (5.5), and the continuity assumed for g, $E(R_n(t, \eta_n, Y_n)) \to 0$ as $n \to \infty$.

Using assumption (5.2) and $|W_{ij}| \leq 1$, a standard argument shows that $var(R_n(t, \eta_n, Y_n)) \to 0$ as $n \to \infty$. With the mean and variance tending to zero the lemma follows.

This lemma shows that $\sqrt{n}\, H_n(t/\sqrt{n}, \eta_n, Y_n)/2t$ converges in probability to γ but the result is not strong enough since an appropriate t must be chosen to fit the data and the nuisance location parameter η_n must be estimated. The following theorem shows that the convergence in the

lemma is uniform and this will allow the construction of a suitable estimate of γ.

<u>Theorem 5.1</u>. Let t_0 and Δ be positive numbers. Then

$$\sup_{\eta_n \in M_n,\ 0 \le t \le t_0} |R_n(t, \eta_n, Y_n)| \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

<u>Proof of Theorem 5.1</u>

This proof is somewhat long and it will be divided into pieces with some lemmas. Begin by choosing any $\varepsilon > 0$. Choose $\delta > 0$ sufficiently small so that $4\sqrt{2p}\, g(0)B_0\, \delta < \varepsilon/6$, where B_0 is specified in (5.2). For this δ form the coverings M_{kn}, $k = 1,\ldots,K$ and the extreme differences d^U_{ijkn} and d^L_{ijkn} of (5.6). Choose $\delta' > 0$ so that $6g(0)\delta' < \varepsilon/6$. Partition the interval $T = [0, t_0]$ into subintervals $T_1,\ldots,T_M$ of width less than δ'. For each $m = 1,\ldots,M$ let

$$t^U_m = \sup\{t: t \in T_m\} \quad \text{and} \quad t^L_m = \inf\{t: t \in T_m\}.$$

Then for each $m = 1,\ldots,M$

$$|t^U_m - t^L_m| < \delta'. \tag{5.7}$$

For each $k = 1,\ldots,K$ and $m = 1,\ldots,M$, define

$$S_{ij}(k,m) = \begin{cases} 1, & \text{if } Y_j - Y_i \in (d^L_{ijkn} - t^U_m/\sqrt{n},\ d^U_{ijkn} - t^L_m/\sqrt{n}) \\ & \text{or } Y_j - Y_i \in (d^L_{ijk} + t^L_m/\sqrt{n},\ d^U_{ijk} + t^U_m/\sqrt{n}) \\ 0, & \text{otherwise} \end{cases}$$

for $1 \le i < j \le n$ and $Q_{k,m} = \sqrt{n} \sum_{i<j} a_{ij} S_{ij}(k,m)$.

<u>Lemma 5.3</u> For each $k = 1,\ldots,K$ and $m = 1,\ldots,M$

$$\sup_{\substack{t,t' \in T_m \\ \eta,\eta' \in M_{kn}}} |R_n(t, \eta, Y) - R_n(t', \eta', Y)| \le Q_{k,m} + 2\gamma\delta'$$

Proof: Fix k and m and write

$$R_n(t, \eta, Y) - R_n(t', \eta', Y) = \sqrt{n} \sum_{i<j} a_{ij}(W_{ij} - W'_{ij}) - 2\gamma(t - t')$$

where $W_{ij} = \phi(|Y_j - Y_i - d_{ij}|, t/\sqrt{n})$, $W'_{ij} = \phi(|Y_j - Y_i - d'_{ij}|, t'/\sqrt{n})$, $d_{ij} = \eta_{jn} - \eta_{in}$ and $d'_{ij} = \eta'_{jn} - \eta'_{in}$. Inspection of possible cases shows that $|W_{ij} - W'_{ij}| \leq S_{ij}(k,m)$ for all $1 \leq i < j \leq n$, all $\eta, \eta' \in M_{kn}$ and all $t, t' \in T_m$. Then the lemma follows.

Since g is continuous at 0, there is a $\delta'' > 0$ such that $|d| < \delta''$ implies $g(d) < 2\, g(0)$. Then from (5.5) there exists an integer N_1 such that

$$(-\delta''/2) \leq d^L_{ijkn} \leq d^U_{ijkn} \leq (\delta''/2) \tag{5.8}$$

for all $k = 1,\dots,K$, all $1 \leq i < j \leq n$ and all $n \geq N_1$

Lemma 5.4. For all $k = 1,\dots,K$ and $m = 1,\dots,M$,

(a) there exists an integer N_2 such that

$$E(Q_{k,m}) \leq (\varepsilon/6) + 4\, g(0)\delta' \quad \text{for all } n \geq N_2, \text{ and}$$

(b) $$\operatorname{var}(Q_{k,m}) \to 0 \text{ as } n \to \infty.$$

Proof:

$$E(Q_{k,m}) = \sqrt{n} \sum_{i<j} a_{ij}\ E(S_{ij}(k,m))$$

$$= \sqrt{n} \sum_{i<j} a_{ij}\{[G(d^U_{ijkn} - t^L_m/\sqrt{n}) - G(d^L_{ijkn} - t^U_m/\sqrt{n})]$$

$$+ [G(d^U_{ijkn} + t^U_m/\sqrt{n}) - G(d^L_{ijkn} + t^L_m/\sqrt{n})]\}$$

$$= \sqrt{n} \sum_{i<j} a_{ij}\{g(\xi_{ij}(k,m))[d^U_{ijkn} - d^L_{ijkn} + (t^U_m - t^L_m)/\sqrt{n}]$$

$$+ g(\nu_{ij}\ (k,m))[d^U_{ijkn} - d^L_{ijkn} + (t^U_m - t^L_m)/\sqrt{n}]\}$$

where $d^L_{ijkn} - t^U_m/\sqrt{n} \le \xi_{ij}(k,m) \le d^U_{ijkn} - t^L_m/\sqrt{n}$ and $d^L_{ijkn} + t^L_m/\sqrt{n} \le$ $\nu_{ij}(k,m) \le d^U_{ijkn} + t^U_m/\sqrt{n}$ for all $1 \le i < j \le n$.

Now there is an integer $N_2 > N_1$ such that $t_0/\sqrt{n} < \delta''/2$ for all $n \ge N_2$. This along with (5.8) implies

$$-\delta'' < \xi_{ij}(k,m) < \delta'' \quad \text{and} \quad -\delta'' < \nu_{ij}(k,m) < \delta''$$

for all $1 \le i < j \le n$ and all $n \ge N_2$. Then by the choice of δ'' it follows that

$$E(Q_{k,m}) \le \sqrt{n}\ 4\ g(0) \sum_{i<j} a_{ij}[d^U_{ijkn} - d^L_{ijkn} + (t^U_m - t^L_m)/\sqrt{n}]$$
$$\le 4\ g(0)\ \sqrt{n} \sum_{i<j} a_{ij}[d^U_{ijkn} - d^L_{ijkn}] + 4\ g(0)\delta'$$

from (5.7) for all $n \ge N_2$. Applying the Cauchy-Schwarz inequality and Lemma 5.1 and 5.2 yields part (a).

The second part of this lemma follows from Lemma 5.1 and the fact that $|S_{ij}(k,m)| \le 1$ with a standard variance argument.

Choose an arbitrary $\varepsilon' > 0$. For each $k = 1, \ldots, K$ and $m = 1, \ldots, M$ choose a sequence $\eta^k_n \in M_{kn}$, $n = 1, 2, \ldots$ and a number $t_m \in T_m$.

<u>Lemma 5.5</u>. There exists an integer N_3 such that

$$P(|R_n(t_m, \eta_n^k, Y_n)| > \varepsilon/3) \leq \varepsilon'/2KM$$

for all $k = 1, \ldots, K$, all $m = 1, \ldots, M$ and all $n \geq N_3$.

<u>Proof</u>: Apply Lemma 5.2 for each of the finite number of pairs (k,m).

With these preliminaries completed, the proof of the theorem will now be completed. For each $k = 1, \ldots, K$ and $m = 1, \ldots, M$ write

$$P(\sup_{t\in T_m, \eta_n \in M_{kn}} |R_n(t, \eta_n, Y_n)| \geq \varepsilon)$$

$$\leq P(\sup_{t\in T_m, \eta_n \in M_{kn}} |R_n(t, \eta_n, Y_n) - R_n(t_m, \eta_n^k, Y_n)| + |R_n(t_m, \eta_n^k, Y_n)| \geq \varepsilon)$$

$$\leq P(Q_{k,m} + 2\gamma\delta' + |R_n(t_m, \eta_n^k, Y_n)| \geq \varepsilon)$$

$$= P(Q_{k,m} - E(Q_{k,m}) + E(Q_{k,m}) + 2\gamma\delta' + R_n(t_m, \eta_n^k, Y_n)| \geq \varepsilon)$$

$$\leq P(Q_{k,m} - E(Q_{k,m}) + |R_n(t_m, \eta_n^k, Y_n)| \geq 2\varepsilon/3)$$

$$\leq P(Q_{k,m} - E(Q_{k,m}) \geq \varepsilon/3) + P(|R_n(t_m, \eta_n^k, Y_n)| \geq \varepsilon/3)$$

$$\leq (9/\varepsilon^2)\,\mathrm{Var}(Q_{k,m}) + (\varepsilon'/2KM),$$

$$\leq \varepsilon'/KM$$

where Lemma 5.3 was used in the third line; Lemma 5.4 (a) and the choice of δ' were used in the fifth line; Chebyshev's inequality and Lemma 5.5 were used in the seventh line; and the last line follows from Lemma 5.4 (b) since the first term in the seventh line above can be made less than $\varepsilon'/2KM$ for

all k,m for n sufficiently large, say for $n \geq N_4 \geq N_3$. Finally,

$$P\left(\sup_{0 \leq t \leq t_0,\, \eta_n \in M_n} |R_n(t, \eta_n, Y_n)| \geq \varepsilon\right)$$

$$\leq \sum_{m=1}^{M} \sum_{k=1}^{K} P\left(\sup_{t \in T_m,\, \eta_n \in M_{kn}} |R_n(t, \eta_n, Y_n)| \geq \varepsilon\right)$$

$$\leq \sum_{m=1}^{M} \sum_{k=1}^{K} \varepsilon'/KM = \varepsilon'$$

for all $n \geq N_4$ and this completes the proof.

This leads to our final result for the parameter τ. Define $\hat{\tau}_{n\alpha} = (\sqrt{12}\,\hat{\gamma}_{n\alpha})^{-1}$ where $\hat{\gamma}_{n\alpha}$ is defined in (5.3).

Theorem 5.2. Under the assumptions (A.1), (A.2), and (B.1), $\hat{\tau}_{n\alpha} \xrightarrow{P} \tau$, for $0 < \alpha < 1$.

Proof: Assuming condition (A.2) holds, first note that (A.2) is equivalent to the condition,

$$\lim_{n \to \infty} \max_{1 \leq i \leq n} |x_{ik}| / \sqrt{n} = 0, \quad \text{for} \quad k = 1, \ldots, p,$$

where x_{ik} is the (i,k)th element of any basis matrix X for Ω_n. It then follows from Sievers (1983) that $||\hat{\mu}_n||$ is bounded in probability where $\hat{\mu}_n \in \Omega_n$ is the best R-predictor of Y_n. This can be used to select a $\Delta > 0$ so that $P(||\hat{\mu}_n|| > \Delta)$ is arbitrarily small for sufficiently large n. From Sievers (1982), $\hat{t}_{n\alpha} \xrightarrow{P} t_\alpha$ as $n \to \infty$ where t_α is a positive constant. Let t_0 of Theorem 5.1 be a number larger than t_α. Then by Theorem 5.1, $|R_n(\hat{t}_{n\alpha}, \hat{\mu}_n, Y_n)| \to 0$ in probability and the conclusion follows.

Theorem 3.1 is the special case of Theorem 5.2 when $a_{ij} = \binom{n}{2}^{-1}$.

References

Bickel, P.J. (1976). Another look at robustness: A review of reviews and some new developments (Reply to Discussant). Scand. J. of Statist.. 3, 167.

Draper, D. (1981). Rank-Based Robust Analysis of Linear Models. Ph.D. thesis, University of California, Berkeley.

Gross, A.M. (1979). Programs for Digital Signal Processing, Digital Processing Committee, IEEE Acoustics, Speech, and Signal Processing Society, IEEE Press. P.S.-11, S-12.

Hettmansperger, T.P. and McKean, J.W. (1978). Statistical inference based on ranks. Psychometrika. 43, 69-79.

Hettmansperger, T.P. and McKean, J.W. (1983). A geometrical interpretation of inferences based on ranks in the linear model. J. Am. Statist. Assoc., To appear.

Huber, P.J. (1973). Robust regression: asymptotics, conjectures, and Monte Carlo. Ann. Statist. 1, p. 799-821.

Jaeckel, L.A. (1972). Estimating regression coefficients by minimizing the dispersion of the residuals. Ann. Math. Statist. 43, 1449-1458.

Jureckova, J. (1971). Nonparametric estimate of regression coefficients. Ann. Math. Statist. 42, 1328-1338.

Kalbfleisch, J.D. and Prentice, R.L. (1980). The Statistical Analysis of Failure Time Data. New York: John Wiley and Sons.

Lehmann, E.L. (1963). Nonparametric confidence intervals for a shift parameter. Ann. Math. Statist. 34, 1507-1512.

Marsaglia, G. and Bray, T.A. (1964). A convenient method for generating normal variables. SIAM Review 6, 260-264.

McKean, J.W. and Hettmansperger, T.P. (1976). Tests of hypothesis in the general linear model based on ranks. Comm. Statist. A5, 693-709.

McKean, J.W. and Hettmansperger, T.P. (1978b). A robust analysis of the general linear model based on one step R-estimates. Biometrika. 65, 571-579.

McKean, J.W. and Ryan, T.A., Jr. (1977). An algorithm for obtaining confidence intervals and point estimates based on ranks in the two sample location problem. Trans. Math. Software 3, 183-185.

McKean, J.W. and Schrader, R.M. (1980). The geometry of robust procedures in linear models. J. R. Statist. Soc.B 42, 366-371.

Miller, R.G. (1981). Survivor Analysis, New York: John Wiley and Sons.

Schweder, T. (1975). Window estimation of the asymptotic variance of rank estimators of location. Scand. J. Statist. 2, 113-126.

Sievers, G.L. (1982). A consistent estimate of a nonparametric scale parameter. Institute of Statistics Mimeo Series #1501, Department of Statistics, Chapel Hill, North Carolina.

Sievers, G.L. (1983). A weighted dispersion function for estimation in linear models. Comm. Statist. 12, 1161-1179.

TABLE 4.1 EMPIRICAL 05 LEVELS

A. UNBALANCED 2 x 2 DESIGN

ERROR DISTRIBUTION

	NORMAL HYPOTH			CN HYPOTH			SCN HYPOTH			LOGNORMAL HYPOTH		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
LS	044	035$^-$	059	026$^-$	035	041	056	051	042	045	066$^+$	039
D_0	038	036	056	044	046	042	050	054	050	034$^-$	055	042
D(.8)	042	036	063	041	042	040	049	059	051	050	075$^+$	054
D(.9)	039	035$^-$	060	032$^-$	035$^-$	029$^-$	047	057	048	040	065$^+$	042
B(.8)	046	030$^-$	064	042	043	044	052	065$^+$	061	057	087$^+$	061
B(.9)	044	032$^-$	058	026$^-$	032$^-$	028$^-$	046	054	052	038	065$^+$	043

B. SMALL PARALLEL DESIGN

	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
LS	057	044	053	024$^-$	038	027$^-$	039	040	031$^-$	044	047	038
D_0	053	038	045	040	038	029$^-$	041	043	038	048	040	042
D(.8)	056	046	045	036	038	028$^-$	040	047	041	060	045	051
D(.9)	051	040	043	028$^-$	024$^-$	019$^-$	039	036	034$^-$	046	034$^-$	033$^-$
B(.8)	056	048	059	033$^-$	040	041	047	052	050	067$^+$	050	073$^+$
B(.9)	050	040	048	025$^-$	026$^-$	024$^-$	033$^-$	035$^-$	027$^-$	044	030$^-$	046

C. LARGE PARALLEL DESIGN

	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
LS	053	043	057	033$^-$	033$^-$	029$^-$	045	046	045	036	045	036
D_0	054	044	056	050	048	046	047	048	040	030$^-$	039	027$^-$
D(.8)	051	047	056	048	046	044	048	050	042	050	064	049
D(.9)	053	045	053	039	046	035$^-$	042	046	040	042	046	040
B(.8)	057	049	063	047	047	045	044	049	046	055	065$^+$	062
B(.9)	052	049	059	032$^-$	039	028$^-$	038	043	037	037	047	045

Table 4.2 EMPIRICAL POWERS: Normal Distribution

A. UNBALANCED 2 x 2 DESIGN

	H_{01}				H_{02}				H_{03}			
λ	0.0	1.3	3.0	6.9	0.0	1.3	3.0	6.9	0.0	1.7	3.4	10.
LS	044	346	642	950	035	335	645	951	059	421	695	985
D_0	038	324	608	926	036	315	599	939	056	390	655	979
D(.8)	042	324	600	924	036	320	601	938	063	383	657	978
B(.8)	046	322	591	913	030	318	583	931	064	376	641	976

B. SMALL PARALLEL DESIGN

	H_{01}				H_{02}				H_{03}			
λ	0.0	2.0	3.2	7.3	0.0	1.9	3.3	7.3	0.0	.9	1.6	3.4
LS	057	464	650	938	044	466	666	947	053	180	304	550
D_0	053	426	615	923	038	418	614	925	045	157	241	471
D(.8)	056	424	622	922	046	412	616	922	045	155	252	483
B(.8)	056	425	602	912	048	402	608	922	059	159	252	497

C. LARGE PARALLEL DESIGN

	H_{01}				H_{02}				H_{03}			
λ	0.0	2.5	5.0	7.2	0.0	2.8	6.2	10.1	0.0	5.5	8.8	14.0
LS	053	562	854	951	043	655	933	994	057	823	944	995
D_0	054	533	835	933	044	621	919	989	056	790	933	993
D(.8)	051	531	835	937	047	623	918	989	056	783	934	991
B(.8)	057	530	832	935	049	607	912	987	063	773	936	989

TABLE 4.3 EMPIRICAL POWERS: Contaminated Normal Distribution

A. 2 x 2 UNBALANCED DESIGN

	H_{01}				H_{02}				H_{03}			
λ	0.0	2.3	6.9	12.3	0.0	2.3	6.9	12.3	0.0	3.4	10.0	28
LS	026	183	350	497	035	186	339	478	041	221	423	707
D_0	044	392	699	863	046	382	677	865	042	422	804	976
D(.8)	041	387	682	858	042	368	663	852	040	414	793	972
B(.8)	042	364	664	858	043	353	661	862	044	403	799	978

B. SMALL PARALLEL DESIGN

	H_{01}				H_{02}				H_{03}			
λ	0.0	2.0	3.2	7.3	0.0	1.9	3.3	7.3	0.0	2.3	4.8	8.2
LS	024	152	222	391	038	156	228	393	027	114	219	327
D_0	040	243	369	673	038	264	388	666	029	192	370	584
D(.8)	036	227	367	670	038	255	377	653	028	181	353	580
B(.8)	033	220	352	663	040	242	362	640	041	181	358	581

C. LARGE PARALLEL DESIGN

	H_{01}				H_{02}				H_{03}			
λ	0.0	2.5	5.0	7.2	0.0	2.8	6.2	10.1	0.0	5.5	8.8	14.0
LS	033	137	246	310	033	178	320	435	029	208	291	436
D_0	050	308	551	734	048	396	687	875	046	503	705	865
D(.8)	048	298	552	730	046	387	686	871	044	501	710	865
B(.8)	047	295	544	732	039	341	661	845	045	490	716	876

TABLE 4.4 EMPIRICAL POWERS: Skewed Contaminated Normal Distribution

A. 2 x 2 UNBALANCED DESIGN

	H_{01}				H_{02}				H_{03}			
λ	0.0	3.0	6.9	12.	0.0	3.0	6.9	12.	0.0	1.7	3.4	10.
LS	056	299	536	739	051	296	545	759	042	181	322	637
D_0	050	385	726	882	054	405	726	906	050	256	440	818
D(.8)	049	400	733	899	059	411	733	914	051	266	454	837
B(.8)	052	392	722	896	065	405	729	918	061	267	437	832

B. SMALL PARALLEL DESIGN

	H_{01}				H_{02}				H_{03}			
λ	0.0	2.0	3.2	7.3	0.0	1.9	3.3	7.3	0.0	2.3	4.8	8.2
LS	039	208	306	532	040	208	312	556	031	171	310	460
D_0	041	263	388	678	043	277	402	711	038	217	403	598
D(.8)	040	276	397	692	047	272	404	708	041	214	404	616
B(.8)	047	266	389	689	052	259	401	693	050	215	408	609

C. LARGE PARALLEL DESIGN

	H_{01}				H_{02}				H_{03}			
λ	0.0	2.5	5.0	7.2	0.0	2.8	6.2	10.1	0.0	5.5	8.8	14.0
LS	045	227	411	535	046	274	481	671	045	338	491	689
D_0	047	363	619	762	048	396	705	885	040	540	720	898
D(.8)	048	376	622	768	050	400	715	883	042	558	731	902
B(.8)	044	367	621	767	049	396	709	885	046	562	733	900

TABLE 4.5 EMPIRICAL POWERS: Lognormal Distribution

A. 2 x 2 UNBALANCED DESIGN

	H_{01}				H_{02}				H_{03}			
λ	0.0	1.3	3.0	6.9	0.0	1.3	3.0	6.9	0.0	1.7	3.4	10.
LS	045	167	335	543	066	196	329	583	039	234	381	695
D_0	034	260	516	807	055	274	507	826	042	365	602	924
D(.8)	050	371	625	867	075	371	638	882	054	480	716	957
B(.8)	057	353	637	882	087	352	633	894	061	482	713	964

B. SMALL PARALLEL DESIGN

	H_{01}				H_{02}				H_{03}			
λ	0.0	2.0	3.2	7.3	0.0	1.9	3.3	7.3	0.0	.9	1.6	3.4
LS	044	236	338	582	047	240	367	588	038	090	143	274
D_0	048	345	480	775	040	330	501	784	042	120	183	369
D(.8)	060	430	593	848	045	430	615	853	051	152	239	475
B(.8)	067	419	590	850	050	408	612	867	073	173	244	475

C. LARGE PARALLEL DESIGN

	H_{01}				H_{02}				H_{03}			
λ	0.0	2.5	5.0	7.2	0.0	2.8	6.2	10.1	0.0	5.5	8.8	14.0
LS	036	265	441	559	045	280	485	655	036	374	521	689
D_0	030	470	753	888	039	497	828	938	027	664	836	953
D(.8)	050	613	850	937	064	649	895	969	049	812	921	976
B(.8)	055	612	870	949	065	658	906	980	062	832	938	991

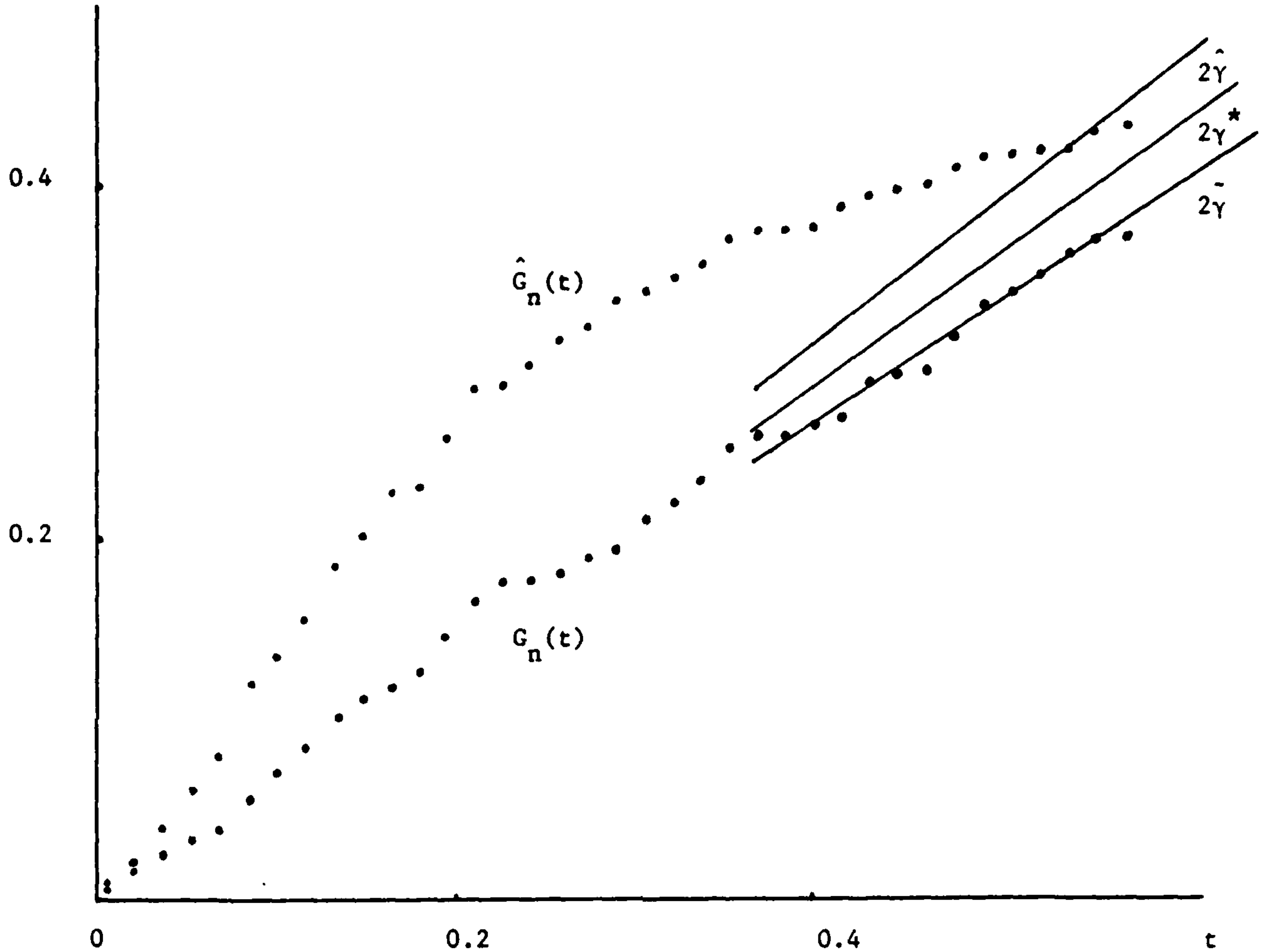


Figure 4.1. $\hat{G}_n(t)$ and $G_n(t)$ are the empirical cumulative distribution functions of the absolute differences of R-residuals and random errors, respectively, from a simulation of lognormal variables with design A. For the slopes, $\hat{\gamma}$ is from (3.5) with $\alpha = .80$, $\tilde{\gamma}$ is similar but based on the random errors and γ^* corrects the $\hat{\gamma}$ as in (4.1).

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report #70	2. GOVT ACCESSION NO. A135792	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) On the Robust Rank Analysis of Linear Models with Nonsymmetric Error Distributions		5. TYPE OF REPORT & PERIOD COVERED Technical Report 1983
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Gerald L. Sievers and Joseph W. McKean		8. CONTRACT OR GRANT NUMBER(s) N00014-78-C-0637
9. PERFORMING ORGANIZATION NAME AND ADDRESS Western Michigan University Kalamazoo, Michigan 49008		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-407
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics and Probability Program		12. REPORT DATE September 1983
		13. NUMBER OF PAGES 31
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

general linear hypothesis; linear model; R-estimates; rank statistics; regression; robust analysis.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Please see next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT

The robust analysis of linear models based on R-estimates involves an estimate of a scale parameter which is used in the analysis as a standardizing constant. The consistency of previous estimates of this scale parameter required that the underlying errors be symmetrically distributed. This assumption is not always warranted, for instance in survival models. A new estimate is proposed for the scale parameter and it is shown to be consistent for nonsymmetric and symmetric error distributions. with this new scale estimate, a complete robust analysis of a linear model can be accomplished without assuming symmetry. The small sample properties of the analysis are examined in a Monte Carlo study of several different situations.

LMED

8